REMARKS

This response responds to the Office Action dated February 21, 2003, in which the Examiner rejected claims 1-5, 13, 16, 19 and 20 under 35 U.S.C. § 102(b) and rejected claims 6-12, 14, 15, 17 and 18 under 35 U.S.C. § 103.

Claim 1 claims a method for extracting information from a natural language text corpus based on a natural language query and claim 16 claims a system for extracting information. The system and method comprise a text analysis unit, a storage means, an indexer, an index, a query manager and a result manager. The text analysis unit is for analyzing a natural language text corpus and a natural language query with respect to surface structure of word tokens and surface syntactic roles of constituents. The storage means is operatively connected to the text analysis unit, and is for storing the analyzed natural language text corpus. The indexer is operatively connected to the storage means and is for indexing the analyzed natural language text corpus. The index is operatively connected to the indexer and is for storing the indexed analyzed natural language text corpus. The query manager is operatively connected to the text analysis unit, and comprises means for creating surface variants of the natural language query, the surface variants being equivalent to the natural language query with respect to lexical meaning of word tokens and surface syntactic roles of constituents, and means for comparing the surface variants and the analyzed natural language query with the indexed analyzed natural language text corpus in the index. The result manager is operatively connected to the index and is for extracting, from the indexed and stored analyzed natural language text corpus,

each portion of text comprising a string of word tokens that matches any one of the surface variants or the analyzed natural language query.

Through the method and structure of the claimed invention creating one or more surface variants of the analyzed natural language query, where the one or more surface variants are equivalent to the natural language query with respect to lexical meaning of word tokens and surface syntactic roles of constituents, as claimed in claims 1 and 16, the claimed invention provides a method and system of extracting information in which the number of matches is increased relative to what it would be if the matches were only verbatim searched. Since the form of the variants where the lexical meaning of word tokens and surface syntactic roles of constituents are preserved in relation to the original query, the matching portion in the text corpus will have the same meaning as the original query. Hence, an increased number of hits will consist of relevant hits. Since the search is performed for each variant, the analysis which needs to be done in real time will be less than when the search is performed using matching patterns so that the method according to the invention as claimed is faster than the method in which matching is done using a regular expression. The prior art does not show, teach or suggest the invention as claimed in claims 1 and 16.

Claims 1-5, 13, 16, 19 and 20 were rejected under 35 U.S.C. § 102(b) as being anticipated by *Julliard* (European Patent Application No. 0 886 226).

Applicants respectfully traverse the Examiner's rejection of the claims under 35 U.S.C. § 102(b). The claims have been reviewed in light of the Office Action, and for

reasons which will be set forth below, it is respectfully requested that the Examiner

withdraw the rejection to the claims and allows the claims to issue.

Julliard appears to disclose a method of searching for information in a text database,

comprising: (a) receiving at least one user input, the user input(s) defining a natural

language expression including one or more words, (b) converting the natural language

expression to a tagged form of the expression, the tagged form including said one or more

words and, for the or each word, a part-of-speech tag associated therewith, (c) applying to

the tagged form one or more grammar rules of the language of the natural language

expression, to derive a regular expression, and (d) analyzing the text database to determine

whether there is a match between said regular expression and a portion of said text

database. (col. 1, lines 31-44) The linguistic search techniques provide a new way to

search for information in a text database. They enable users to find portions of a text which

match multiword expressions given by the user. Matches include possible variations that

are relevant with the initial criteria from a linguistic point of view including simple

inflections like plural/singular, masculine/feminine or conjugated verbs and even more

complex variations like the insertion of additional adjectives, adverbs, etc. in between the

words specified by the user. This technique can complement conventional full text search

engines by reducing the number of retrieved documents that are inconsistent with the query.

(col. 2, lines 32-44)

Thus, *Julliard* merely discloses converting the natural language expression to a

tagged form of the expression, where the tagged form includes one or more words for each

word, such as plural/singular, masculine/feminine or conjugated verbs. However, nothing in *Julliard* shows, teaches or suggest creating one or more surface variants which are <u>equivalent to the natural language query with respect to lexical meaning of word tokens and surface syntactic roles of constituents</u> as claimed in claims 1 and 16. Rather, *Julliard* merely discloses adding part of speech tags to each word of a query and deriving a regular expression from the query by means of grammatical rules. The regular expression in *Julliard* indicates that in addition to the words of the query, a natural language expression of a text which is searched may include additional words of certain types in connection to the words of the query and still constitute a match. Thus, *Julliard* does not show, teach or suggest syntactic roles of word tokens of the regular expression with respect to the query as claimed in claims 1 and 16.

Since nothing in *Julliard* shows, teaches or suggests creating one or more surface variants which are <u>equivalent with respect to lexical meaning of word tokens and surface syntactic roles of constituents</u> as claimed in claims 1 and 16, it is respectfully requested that the Examiner withdraw the rejection to claims 1 and 16 under 35 U.S.C. § 102(b).

Claims 2-5, 13, 19 and 20 depend from claims 1 and 16 and recite additional features. It is respectfully submitted that claims 2-5, 13, 19 and 20 would not have been anticipated by *Julliard* within the meaning of 35 U.S.C. § 102(b) at least for the reasons as set forth above. Thus, it is respectfully requested that the Examiner withdraw the rejection to claims 2-5, 13, 19 and 20 under 35 U.S.C. § 102(b).

Claims 6-12 and 18 were rejected under 35 U.S.C. § 103 as being anticipated over *Julliard* in view of *Arampatzis et al.* ("An Evaluation of Linguistically-motivated Indexing Schemes," hereinafter referred to as *Arampatzis-Indexing*). Claims 14, 15 and 17 were rejected under 35 U.S.C. § 103 as being unpatentable over *Julliard* in view of *Arampatzis et al.* ("Linguistically-motivated Information Retrieval," hereinafter referred to as *Arampatzis-Retrieval*).

Applicants respectfully traverse the Examiner's rejection of the claims under 35 U.S.C. § 103. The claims have been reviewed in light of the Office Action, and for reasons which will be set forth below, it is respectfully requested that the Examiner withdraw the rejection to the claims, and allows the claims to issue.

As indicated above, since nothing in *Julliard* shows, teaches or suggest the primary feature as claimed in claims 1 and 16, it is respectfully requested that the combination with the secondary references to *Arampatzis-Indexing* and *Arampatzis-Retrieval* would not overcome the deficiencies of the primary reference. Furthermore, neither *Arampatzis-Indexing* or *Arampatzis-Retrieval* shows, teaches or suggests that any variants of a query should be created, which variants are equivalent to a query with respect to the lexical meanings of word tokens and surface syntactic roles of constituents. Therefore, it is respectfully submitted that the combination of the primary reference and the secondary references will not overcome the deficiencies of the primary reference. It is respectfully requested that the Examiner withdraw the rejection to claims 6-12, 14, 15, 17 and 18 under 35 U.S.C. § 103.

(05/03)

The prior art of record, which is not relied upon, is acknowledged. The references taken singularly or in combination do not anticipate or make obvious the claimed invention.

Thus, it now appears that the application is in condition for reconsideration and allowance. Reconsideration and allowance at an earlier date are respectfully requested.

If for any reason the Examiner feels that the application is not now in condition for allowance, it is respectfully requested that the Examiner contact, by telephone, the applicants' undersigned attorney at the indicated telephone number to arrange for an interview to expedite the disposition of this case.

In the event that this paper is not timely filed within the currently set shortened statutory period, applicants respectfully petition for an appropriate extension of time. The fees for such extension of time may be charged to our Deposit Account No. 02-4800.

In the event that any additional fees are due with this paper, please charge our Deposit Account No. 02-4800.

Respectfully submitted,

BURNS, DOANE, SWECKER & MATHIS, L.L.P.

Date: May 21, 2003   By:

Ellen Marcie Emas
Registration No. 32,131

P.O. Box 1404
Alexandria, Virginia 22313-1404
(703) 836-6620

(05/03)